The phenomenon of spontaneous replica symmetry breaking in complex statistical mechanics systems

Francesco Guerra

Dipartimento di Fisica, Università di Roma "La Sapienza" and Istituto Nazionale di Fisica Nucleare, Sezione di Roma, Piazzale Aldo Moro 5, 00185 Roma (Italy)

E-mail: francesco.guerra@roma1.infn.it

Abstract. We analyze the main aspects of the phenomenon of spontaneous replica symmetry breaking, introduced by Giorgio Parisi. We work in the frame of real replicas, by taking into account the simple case of the random energy model. In particular, we study the phase space diagram for systems of coupled replicas, and the connected phase transitions. Our considerations can be generalized to the more complicated models of mean field spin glasses and neural networks. We report also about a letter of Ettore Majorana, written in December 1937 to his uncle Dante, very interesting for its methodological content.

The phenomenon of "spontaneous replica symmetry breaking", for some mean field models for spin glasses, as the Sherrington-Kirkpatrick model [1], was discovered by Giorgio Parisi, in the frame of his celebrated "replica trick" (see for example [2]). Quite recently a rigorous treatment has been possible by using interpolation methods [3].

"Spontaneous replica symmetry breaking" can be extended to other cases of interest. Typical cases are:

– Spin glasses with coupled replicas. Here a partial answer has been given by Michel Talagrand [4], in the frame of his treatment of the replica symmetric region and the rigorous establishment of the Parisi solution for the free energy. However, the general structure of spontaneous replica symmetry breaking for coupled replicas has not been achieved yet. This would be very important, also for the property of ultra-metricity of the states.

– Bipartite spin glasses. Here we find that the replica symmetric Ansatz, and the fully broken replica case, require a sort of mini-max principle for the free energy (see for example [5]).

– One can consider also neural networks of the Hopfield type. We modify the interaction, so that the learned words, which act as quenched noise, are not made by a string of ± 1 variables, but by a string of unit Gaussian random variables. This helps in the treatment of interpolation, which would be more cumbersome in the standard case. We see here how the mini-max principle emerges already at the level of the replica symmetric Ansatz. Replica symmetry breaking is introduced in the natural way, and still implies a mini-max structure [6, 7].

– Of course the mini-max structure adds difficulty to the rigorous treatment of these models. The nice interpolation inequalities for the mean field spin glass no longer hold. This gives a real challenge for future research.

Coming from spontaneous replica symmetry breaking, the main features of the Parisi variational principle for the free energy are the following:

– the order parameter is a functional order parameter (while for example in the elementary mean field ferromagnetic model, the Curie-Weiss model, the order parameter is simply a constant, with the meaning of the magnetization of the system). The functional order parameter gives all physical properties of the model: thermodynamic quantities (as free energy per site, internal energy, entropy), the states, the fluctuations.

– the variational principle is expressed through a *sup* for the free energy, and NOT through an *inf*, as it happens for example in the case of the entropy principle of equilibrium statistical mechanics. In the Curie-Weiss model the variational principle is given through an *inf*, so it is a genuine entropy principle.

Therefore, the Parisi variational principle is a completely new structure.

We have studied some of its properties, by finding the Legendre form of the variational principle and its Legendre dual. These results will appear in a paper now in preparation. In this frame, let us only recall that the Legendre dual of the Parisi functional order parameter is connected with the general shape of the interaction.

As we have already remarked, the problem of extending the broken replica structure to more complicated cases has not found a general complete solution yet.

In this paper, we will stress that spontaneous replica symmetry breaking is a real physical phenomenon, with its physical consequences, and with the meaning literally expressed.

This is particularly evident if we work with real replicas, as it is possible to do, and not inside a frame where the number of replicas is allowed to go to zero.

In order to simplify the treatment, we work in a completely elementary frame, that of the so called random energy model, introduced by Derrida some years ago [8], in order to mimic some of the features of spin glasses. Our aims are pedagogical.

Let me recall also that I was told by Giorgio Parisi that a kind of random energy model was considered in the 70-ies by Nicola Cabibbo, who did not publish any result because it was considered too simple. This simplicity will give us an elementary frame in order to discuss the phenomenon of spontaneous replica symmetry breaking.

The configuration space is given by an Ising system on N sites

$$\sigma : (1, 2, \dots, N) \ni i \to \sigma_i = \pm 1.$$

Let us introduce a system of (quenched) random variables J_σ, indexed by the configurations σ, which are independent centered unitary Gaussian, so that we have the averages $\mathbb{E}(J_\sigma) = 0$, and $\mathbb{E}(J_\sigma J_{\sigma'}) = \delta_{\sigma\sigma'}$. The symbol $\delta_{\sigma\sigma'}$ (overlap between the two configurations) means 1 if the two configurations are equal, and zero otherwise. Now the partition random function is given by

$$Z_N(\beta, J) \equiv \sum_\sigma \exp(\beta \sqrt{\frac{N}{2}} J_\sigma),$$

where $\beta \geq 0$ is a parameter with the physical meaning of inverse temperature. Of course, we have the random free energy $F_N(\beta, J)$ related to Z in the known form

$$Z_N(\beta, J) = \exp(-\beta F_N(\beta, J)).$$

Call $\mathbb{E}$ the average with respect to the quenched $J_{..}$ variables. Quenched means that they do not participate to thermal equilibrium, but act as an external random environment.

Of course this is a very simple and completely solvable model. Let us see some results. Firstly, a thermodynamically corrected annealed inequality is easily established.

For any $0 < m \leq 1$, we have for the quenched free energy per site the inequality

$$N^{-1} \mathbb{E} \log \sum_\sigma \exp(\beta \sqrt{\frac{N}{2}} J_\sigma) \leq m^{-1} N^{-1} \mathbb{E} \log \sum_\sigma \exp(m\beta \sqrt{\frac{N}{2}} J_\sigma),$$

which holds for any spin system for purely thermodynamic reasons (positivity of the entropy).

Now we apply the annealed inequality, coming from convexity, $\mathbb{E}\log .. \leq \log \mathbb{E}...$ The $\mathbb{E}$ of the *Boltzmannfaktor* is immediately calculated.

$$\mathbb{E}\exp(\beta\sqrt{\frac{N}{2}}J_\sigma) = \exp(\frac{\beta^2}{4}N).$$

By taking into account that $\sum_\sigma = 2^N$, we end with the inequality

$$N^{-1}\mathbb{E}\log\sum_\sigma \exp(\beta\sqrt{\frac{N}{2}}J_\sigma) \leq \frac{\log 2}{m} + \frac{\beta^2}{4}m.$$

The inequality is made the best possible by taking the minimum with respect to the parameter m, for each value of β. The result is then the following. Define the critical temperature through $\beta_c = 2\sqrt{\log 2}$. Then, for $\beta \leq \beta_c$, the minimum is at $m(\beta) = 1$, and its value is $\log 2 + \beta^2/4$. While for $\beta \geq \beta_c$, the minimum is at $m(\beta) = \beta_c/\beta$, and its value is $\beta\sqrt{\log 2}$. Therefore, the optimal $m(\beta) = \max(1, \beta_c/\beta)$ is continuous in β, but its first derivative is discontinuous at the critical point.

With some little additional effort it can be shown that these bounds, given by the stated variational principle, are in fact the true values, in the thermodynamic limit $N \to \infty$.

Therefore we have

$$\lim_{N\to\infty} N^{-1}\mathbb{E}\log\sum_\sigma \exp(\beta\sqrt{\frac{N}{2}}J_\sigma) \equiv A(\beta),$$

where $A(\beta) = \log 2 + \beta^2/4$, for $\beta \leq \beta_c$, and $A(\beta) = \beta\sqrt{\log 2}$, for $\beta \geq \beta_c$.

Fluctuations are easily controlled, so that the above limit holds also without the $\mathbb{E}$ average, J-almost surely.

Let us show the meaning of the optimal order parameter $m(\beta)$. Introduce, as usual, the random Boltzmann-Gibbs average ω for observables F depending on σ:

$$\omega(F) = Z^{-1}\sum_\sigma F\exp..,$$

where $\exp ..$ is the *Boltzmannfaktor*. By taking s, $s = 1, 2, \ldots$, replicas of the system, indexed by the variables σ_i^a, $a = 1, \ldots, s$, $i = 1, \ldots, N$, we can introduce the product state Ω on observables F explicitly as follows

$$\Omega(F) = Z^{-s}\sum_{\sigma^1\ldots\sigma^s} F\exp(\beta\sqrt{\frac{N}{2}}(J_{\sigma^1} + \cdots + J_{\sigma^s}),$$

where the *Boltzmannfaktor* is the product for each replica.

We introduce also the quenched averages $\langle F\rangle = \mathbb{E}\Omega(F)$. The thermodynamic limits are well defined.

Then, an easy calculation, involving integration by parts on the J variables, shows:

$$\partial_\beta A(\beta) = \frac{\beta}{2}(1 - \langle\delta_{12}\rangle).$$

Here, two replicas are involved σ^1 and σ^2, and δ_{12} is their $0-1$ overlap.

From the explicit expression given above we have also $\langle\delta_{12}\rangle = 0$, below the critical value for β, and $\langle\delta_{12}\rangle = 1 - \beta_c/\beta = 1 - m(\beta)$, above the critical value for β. Therefore, by taking into account that δ_{12} takes the values $(0,1)$, we see that the physical meaning of $m(\beta)$ is simply the probability that δ_{12} takes the value 0 under the global average $\langle..\rangle$.

Analogous considerations can be developed for the calculation of the $\langle..\rangle$ averages for all kinds of overlap products among different replicas. For example we have

$$\langle\delta_{12}\delta_{13}\rangle = \frac{1}{2}\langle\delta_{12}\rangle + \frac{1}{2}\langle\delta_{12}\rangle^2 = 1 - \frac{3}{2}m(\beta) + \frac{1}{2}m(\beta)^2,$$

$$\langle\delta_{12}\delta_{34}\rangle = \frac{1}{3}\langle\delta_{12}\rangle + \frac{2}{3}\langle\delta_{12}\rangle^2 = 1 - \frac{5}{3}m(\beta) + \frac{2}{3}m(\beta)^2,$$

as an elementary instance of the so called Ghirlanda-Guerra identities [9].

Of course, we have complete symmetry among all replicas, so that for example $\langle\delta_{12}\rangle = \langle\delta_{13}\rangle$, $\langle\delta_{12}\delta_{23}\rangle = \langle\delta_{34}\delta_{35}\rangle$, $\langle\delta_{12}\delta_{34}\rangle = \langle\delta_{25}\delta_{36}\rangle$, and so on.

Now we show that the region $\beta > \beta_c$ is really characterized by spontaneous replica symmetry breaking, in the literal sense.

Since here we have a kind of mean field system, replica symmetry breaking should be achieved along these steps: firstly we break explicitly the symmetry according to some small parameter λ, then we go to the limit $N \to \infty$, then we bring $\lambda \to 0$, and see what happens. Therefore, we start from a system made by the σ variables for two replicas, for example σ^1 and σ^2, and introduce, for some $\lambda \geq 0$,

$$\tilde{A}_N(\beta,\lambda) = (2N)^{-1}\mathbb{E}\log\sum_{\sigma^1\sigma^2}\exp(\beta\sqrt{\frac{N}{2}}(J_{\sigma^1}+J_{\sigma^2}))\exp(\frac{1}{2}\lambda N\delta_{\sigma^1\sigma^2}).$$

There is an explicit coupling between the σ^1 and the σ^2. Due to the N factor the coupling does have thermodynamic effects. The replicated systems will contain the couples of replicas $(12), (34), (56), \ldots$. If $\lambda > 0$ replica symmetry is explicitly broken. The surviving symmetries are $(12), (34)$ and so on. But there is no (13) symmetry for example.

However, the model with coupled replicas is also explicitly solvable. In order to understand this, let us write

$$\exp(\frac{1}{2}\lambda N\delta_{\sigma^1\sigma^2}) = (\exp(\frac{1}{2}\lambda N) - 1)\delta_{\sigma^1\sigma^2} + 1 \simeq \exp(\frac{1}{2}\lambda N)\delta_{\sigma^1\sigma^2} + 1,$$

since the -1 term has no thermodynamic effect.

Therefore, by substitution in the previous

$$\tilde{A}_N(\beta,\lambda) \simeq (2N)^{-1}\mathbb{E}\log(Z_N(2\beta)\exp(\frac{1}{2}\lambda N) + Z_N^2(\beta)).$$

Now consider the dominant terms in the thermodynamic limit

$$Z_N(2\beta)\exp(\frac{1}{2}\lambda N) \simeq \exp((A(2\beta) + \frac{1}{2}\lambda)N),$$

$$Z_N^2(\beta) \simeq \exp(2A(\beta)N).$$

As $N \to \infty$, between the two competitors in the sum, the highest will prevail, so that we have

$$\lim_{N\to\infty}\tilde{A}_N(\beta,\lambda) = \tilde{A}(\beta,\lambda) = \max(\frac{1}{2}A(2\beta) + \frac{1}{4}\lambda, A(\beta)).$$

Since $A(2\beta)$ and $A(\beta)$ are explicitly known, as we have shown before, we can easily calculate $\tilde{A}(\beta,\lambda)$. The final result is the following.

There is a critical line for λ, defined by two interlaced parabolas in β, as follows: $\lambda_c(\beta) = 2\log 2 - \beta^2$ for $0 \leq \beta \leq \sqrt{\log 2} = \beta_c/2$, $\lambda_c(\beta) = (\beta - \beta_c)^2$ for $\beta_c/2 \leq \beta \leq \beta_c$, $\lambda_c(\beta) = 0$ for $\beta \geq \beta_c$.

For $\lambda \leq \lambda_c, \beta \leq \beta_c$ we have $\tilde{A}(\beta, \lambda) = A(\beta)$, i.e. the λ coupling has absolutely no effect.

On the other hand, if $\lambda \geq \lambda_c$, then we have:

$$\tilde{A}(\beta, \lambda) = \frac{1}{2}\log 2 + \frac{1}{2}\beta^2 + \frac{1}{4}\lambda,$$

for $0 \leq \beta \leq \beta_c/2$, and

$$\tilde{A}(\beta, \lambda) = \beta\sqrt{\log 2} + \frac{1}{4}\lambda,$$

for $\beta_c/2 \leq \beta \leq \beta_c$. Therefore for $\lambda > \lambda_c$ the coupling is effective and $\tilde{A}(\beta, \lambda)$ does depend on λ. Notice that in any case

$$\lim_{\lambda \to 0} \tilde{A}(\beta, \lambda) = A(\beta).$$

This procedure can be easily generalized to the case where we couple any number of s replicas through parameters $\lambda_{12}, \ldots \lambda_{1s}, \ldots \lambda_{s-1s}$, and let them go to zero according to different patterns, after the thermodynamic limit.

Now comes the interesting surprise. By a direct calculation we have

$$\partial_\lambda \tilde{A}(\beta, \lambda) = \frac{1}{4}\langle \delta_{12} \rangle_{\beta\lambda},$$

$$\partial_\beta \tilde{A}(\beta, \lambda) = \frac{\beta}{2}(1 + \langle \delta_{12} \rangle_{\beta\lambda} - 2\langle \delta_{13} \rangle_{\beta\lambda}),$$

where integration by parts on J is exploited again (but now replicas $(12), (34), \ldots$ are coupled). In the notation $\langle .. \rangle_{\beta\lambda}$ we have emphasized the $\beta\lambda$ dependence of the averages.

Since the explicit expression of $\tilde{A}(\beta, \lambda)$ is know, by making the derivatives, through simple calculations, we get $\langle \delta_{12} \rangle_{\beta\lambda}$ and $\langle \delta_{13} \rangle_{\beta\lambda}$ in the following explicit form:

$$\langle \delta_{12} \rangle_{\beta\lambda} = 0, \langle \delta_{13} \rangle_{\beta\lambda} = 0,$$

for $\lambda < \lambda_c, \beta \leq \beta_c$, as expected. Moreover

$$\langle \delta_{12} \rangle_{\beta\lambda} = 1, \langle \delta_{13} \rangle_{\beta\lambda} = 0,$$

for $\lambda > \lambda_c, 0 \leq \beta \leq \beta_c/2$. Crossing the critical line $\langle \delta_{12} \rangle_{\beta\lambda}$ gets a bang-bang transition, from the minimal value 0 to the maximal value 1. On the other hand, the interaction is not strong enough to displace $\langle \delta_{13} \rangle_{\beta\lambda}$ from its 0 value. However, if $\lambda > \lambda_c, \beta_c/2 \leq \beta \leq \beta_c$, then we have

$$\langle \delta_{12} \rangle_{\beta\lambda} = 1, \quad \langle \delta_{13} \rangle_{\beta\lambda} = 1 - \frac{\beta_c}{2\beta}.$$

Notice the factor 1/2. From these explicit expressions, we have the spontaneous replica symmetry breaking in the most direct form. In fact for $\beta > \beta_c$ we have

$$\langle \delta_{12} \rangle_{\beta 0^+} = \lim_{\lambda \to 0} \langle \delta_{12} \rangle_{\beta\lambda} = 1 > \langle \delta_{12} \rangle_\beta,$$

$$\langle \delta_{13} \rangle_{\beta 0^+} = \lim_{\lambda \to 0} \langle \delta_{13} \rangle_{\beta\lambda} = 1 - \frac{1}{2}\frac{\beta_c}{\beta} > \langle \delta_{12} \rangle_\beta = 1 - \frac{\beta_c}{\beta},$$

where we have indicated with $\langle .. \rangle_\beta$ the averages for uncoupled replicas ($\lambda = 0$ from the beginning) at inverse temperature β. Clearly $\langle \delta_{12} \rangle_{\beta 0^+}$ is different from $\langle \delta_{13} \rangle_{\beta 0^+}$. Replica symmetry is spontaneously broken! The chain $\lambda > 0, N \to \infty, \lambda \to 0$ gives different results from those where $\lambda = 0$ from the beginning.

Of course, we have also to understand how the broken quantities are connected with the original ones, so to recognize the pattern of the spontaneously breaking of the symmetry. Fortunately in this case the problem is very simple. We have already seen the validity of

$$\lim_{\lambda \to 0} \tilde{A}(\beta, \lambda) = A(\beta, \lambda),$$

telling us that the free energy is continuous in λ. But we have more. In fact, we can easily check that we have also for the derivatives

$$\lim_{\lambda \to 0} \partial_\beta \tilde{A}(\beta, \lambda) = \partial_\beta A(\beta).$$

Therefore, we have continuity also for the internal energy and the entropy. Spontaneous replica symmetry breaking is able to change the state at $\lambda = 0^+$, with respect to $\lambda = 0$, but preserves the free energy, the internal energy and the entropy.

We leave to the interested reader the enquire about the fate of the Ghirlanda-Guerra identities for systems with coupled replicas, and for generic spontaneously broken states.

Analogous considerations can be developed for more complicated cases, as the Sherrington-Kirkpatrick model, and neural networks. Here Parisi functional order parameter plays a central role.

Acknowledgments
This work was supported in part by MIUR (Italian Minister of Instruction, University and Research), University of Rome “La Sapienza”, and INFN (Italian National Institute for Nuclear Physics).

Appendix
As a kind of introduction to the topics of complex systems, in a very peculiar historical perspective, we would like to report here about an important letter of Ettore Majorana to his uncle Dante, dated December 27, 1937, relevant for its methodological considerations.

Ettore Majorana (1906-1939), the famous Theoretical Physicists, is worldwide known for his remarkable achievements in the fields of atomic and molecular physics, nuclear physics and elementary particles [10]. His ideas are well up to date, as for example the ongoing search for the Majorana neutrino shows. Moreover, his personal history attracted the interest of many people, because he disappeared suddenly, under still mysterious circumstances, in March 1938, only few months after his direct appointment as Full Professor of Theoretical Physics at the Royal University of Naples.

The letter is written to the uncle Dante, brother of the father Fabio, in December 1937, only one month after the appointment in Naples. A copy can be found at the Majorana Fund of the Regional University Library in Catania, the city in Sicily where the Majorana Family took origin. This is a gigantic Fund, amounting to more than one hundred thousand documents, not fully classified in a complete catalog yet. It is an invaluable source of information about the Majorana Family, starting from the Patriarch Salvatore Majorana Calatabiano (1825-1897), *lawyer, member of the Parliament, minister, senator, professor of the University of Catania, economist, jurist, orator, writer*, and his sons Giuseppe (1863-1940), Angelo (1865-1910), Quirino (1871-1957), Dante (1874-1955), Fabio (1875-1934), the father of Ettore, and the daughters Elvira, Emilia. There are many important documents in this Fund, not all fully explored yet, related to Ettore, and his disappearance, as seen from the Catania observatory of the uncle Giuseppe.

Here is the text of the letter, copied by hand of a member of the Majorana Family in Catania, in our literal English translation.

Roma 27 - 12 - 1937 XVI
Dear uncle Dante,
I thank you very warmly for your greetings and wishes, which are cordially reciprocated to you and your relatives. Thank you also for your comments on the method. Allow me to add an impression of mine. I believe in the unity of science, but just because I believe seriously in it, I think that, until different sciences will practically exist with different objects, no error would be so pernicious as the confusion of methods. In particular, the mathematical method can not be of any substantial utility in sciences that are presently extraneous to physics. In other words, if some day the mathematical face of the simplest facts of life and conscience will be discovered, most certainly this will not happen for a natural evolution of biology and psycology, but only because some additional radical renewal of the general principles of physics will allow to extend its domain in fields, that are still extraneous to it. The most significant example is given by chemistry, that, after a long and very glorious life as an independent science, it has been completely absorbed by physics in the last years. This has been made possible by the rise of quantum mechanics, while no useful connection could be established between chemistry and classical mechanics. While waiting that physics will perform new miracles, it would be right to recommend to the scholars of other disciplines to rely on the methods proper to each discipline, without searching for models or suggestions in today physics, even less in yesterday's. Because the kind of physics, that someday will tell the definite truth on biological or moral facts, is something completely outside of our understanding.
Affectionate greetings
Ettore

These deep and precise methodological considerations are very impressive. In the present times, some typical biological problems are considered in the frame of methods typical of Physics, as for example neural networks, immune systems, animal behavior. This is perhaps an indication that the development, foresighted by Majorana, has taken place, at least in part. The modern Physics of complex systems (see for example [11] for an inspired overview) is perhaps the bridge toward the beginning of the assimilation of Biology into Physics.

References

[1] Sherrington D and Kirkpatrick S 1975 Solvable Model of a Spin-Glass *Phys. Rev. Lett.* **35** 1792
[2] Mézard M, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
[3] Guerra F 2003 Broken Replica Symmetry Bounds in the Mean Field Spin Glass Model *Comm. Math. Phys.* **233** 1
[4] Talagrand M 2003 *Spin Glasses: a Challenge for Mathematicians. Cavity and Mean Field Models* (Berlin: Springer)
[5] Barra A, Genovese G and Guerra F 2011 Equilibrium statistical mechanics of bipartite spin systems *J. Phys. A: Math. Theor.* **44** 245002
[6] Barra A, Genovese G and Guerra F 2010 The replica symmetric behavior of the analogical neural network *J. Stat. Phys.* **140** 784
[7] Barra A, Genovese G, Guerra F and Tantari D 2012 How glassy are neural networks? *J. Stat. Mech.* P07009
[8] Derrida B 1981 Random-Energy Model: an Exactly Solvable Model of Disordered Systems *Phys. Rev.* B **24** 2613
[9] Ghirlanda S and Guerra F 1998 General Properties of Overlap Probability Distributions in Disordered Spin Systems. Towards Parisi Ultrametricity *J. Phys. A: Math. Gen.* **31** 9149
[10] Guerra F and Robotti N 2008 *Ettore Majorana. Aspects of his scientific and academic activity* (Pisa: Edizioni della Scuola Normale Superiore)
[11] Parisi G 1999 Complex Systems: a Physicist's Viewpoint *Physica* A **263** 557